Penalized denoising of vehicle trajectories collected by a swarm of drones

Georg Anagnostopoulos*[1], Emmanouil Barmpounakis[1], and Nikolas Geroliminis[1]

[1]Urban Transport Systems Laboratory, EPFL, Switzerland

SHORT SUMMARY

Vehicle trajectory datasets collected in urban traffic environments with drones pose unique challenges in terms of denoising due to extensive visual restrictions, perspective distortions and human-induced errors. This article taps into the unexplored potential of penalties in the context of vehicle trajectory reconstruction with the example of the massive pNEUMA dataset. We contribute to the literature by shifting the focus of denoising from smoothing to anomaly detection. Specifically, we distinguish between stationary and non-stationary errors and argue that the latter accounts for the largest part of the noise. We propose a re-purposing of the Butterworth filter for the detection of anomalous events and enforce spatial autocorrelation constraints on the errors with functional data analysis. The calibration of our reconstruction makes further use of penalties and is inspired by the theory of human-machine interaction. Our method can be used for quantifying autocorrelated errors or for identifying network segments that are devoid of errors.

Keywords: anomaly detection, Butterworth filter, functional data analysis, penalized denoising.

1. INTRODUCTION

Traffic monitoring of urban environments with drones is challenging in terms of denoising due to extensive visual restrictions, perspective distortions and human-induced errors. The collected vehicle trajectory datasets lack a reliable ground truth, so a penalized (or regularized) approach in denoising becomes especially relevant. However, the potential of penalties in the context of vehicle trajectory reconstruction remains unexplored. As a case study for our reconstruction, we will use the pNEUMA dataset, a collection of naturalistic vehicle trajectories captured by a swarm of drones hovering over the dense urban environment of Athens, Greece. The design of the experiment to collect this massive dataset along with suggested applications was introduced in (Barmpounakis & Geroliminis, 2020).

Despite the fact that a bird-eye view, compared to an angle view, decreases the hidden points between vehicles to a minimum and the advances in computer vision provide opportunities for higher quality of trajectory extraction, vehicle trajectories in the pNEUMA dataset contain errors that require a careful consideration. Many researchers in the field of intelligent transport systems (ITS) have attempted to recover or reconstruct noise-free vehicle trajectories, such as (Montanino & Punzo, 2015; Coifman & Li, 2017; Dong et al., 2021). Simply, there exist two kinds of noise: white noise and anomalies. White noise is assumed to be independently and identically distributed with constant variance and zero mean. The above assumptions do not apply to anomalies that can be described as asymmetric, abrupt events. Anomalies are also different from outliers, extreme but feasible values of an underlying phenomenon, for example a harsh acceleration or a harsh braking event (Vlahogianni & Barmpounakis, 2017). Outliers are valuable and should be preserved along the process. Some authors, such as (Montanino & Punzo, 2015), do not draw this distinction and refer to all undesirable values as outliers.

2. METHODOLOGY

Our methodology builds on the distinction between stationary and non-stationary, autocorrelated errors and consists of 5 parts: filter design, functional data analysis of the non-stationary errors, theory-driven validation in absence of a ground truth (penalizing false negatives), calibration with penalties (penalizing false positives) and formulation of the objective function for determining the optimal parameters of our reconstruction.

Design of the Butterworth filter

Anomaly detection typically involves using some kind of filter. A classical choice in the field of digital signal processing (DSP), is the Butterworth filter. Surprisingly, despite the excellent denoising properties of this filter, in the literature of vehicle trajectory reconstruction it is used just as a smoothing technique (Montanino & Punzo, 2015; Dong et al., 2021), not for anomaly detection. Specifically, (Montanino & Punzo, 2015) remove outliers by imposing predetermined thresholds on longitudinal acceleration and (Dong et al., 2021) detect anomalies with a more involved rule-based system.

More formally, the magnitude squared frequency response of an Nth order Butterworth filter with cut-off frequency Ω_c is given by the following equation (Taylor, 2011):

$$|H(\Omega)|^2 = \frac{1}{1+\varepsilon^2(\Omega/\Omega_c)^{2N}}, \tag{1}$$

where Ω is the frequency and $H(\Omega)$ denotes the transfer function. The parameter ε controls the magnitude at the critical frequency Ω_c. Usually, we set $\varepsilon = 1$. The order N represents the aggressiveness of the filter. In the limit as N approaches infinity, we obtain the ideal *"brick wall response"*. In practice, low orders are used. For example (Montanino & Punzo, 2015) set $N = 1$ and (Dong et al., 2021) set $N = 2$. A very nice property of this filter is having *"maximally flat magnitude"* in the pass band (Haslwanter, 2021). In contrast to moving averages, which tend to oscillate in the low frequencies, the Butterworth filter suppresses only high frequencies, Figure 1.

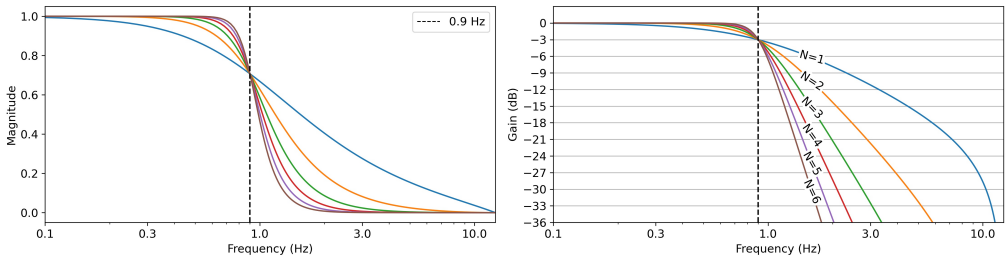


Figure 1: Butterworth filter specification.

After different filter parameterizations, here we will demonstrate how a single Butterworth filter specification with $N = 2$ and $\Omega_c = 0.9$ Hz can be used for detecting anomalies in two trajectory metrics: speed v and azimuth θ, which are noisy realizations of an unobserved ground truth. As one might suspect, there is a catch: we are not primarily interested in the filtered metrics $\widetilde{v}, \widetilde{\theta}$ as such, but in the absolute differences $|v-\widetilde{v}|, |\theta-\widetilde{\theta}|$. For ease of notation and regardless of the metric considered, we will refer to such absolute differences as some error function y, which is specific to each vehicle. At its current form, y cannot be used for anomaly detection. Now the problem is that, strictly speaking, y is not even a function, but rather a set of (noisy) observations y_j at n consecutive timestamps t_j. Therefore, our objective is to find a smooth analytical expression for y in continuous time.

Functional data analysis of non-stationary errors

The motivation of having a smooth error function comes from the assumption that neighboring anomalous observations should be strongly and positively autocorrelated. To reiterate, anomalies are not anything like white noise: they are *"non-stationary, autocorrelated errors"* (J. O. Ramsay & Silverman, 2005). Of course, smoothing of the errors does not need to be an analytical fit and can be also achieved with Gaussian or exponential kernels. The power of the analytical approach is that it maintains most of the original drift without overfitting the data. This process can be fully automatic without depending on sensitive parameters such as various bandwidths. On the other hand, fitting non-stationary error functions can be challenging. Fortunately, there is a whole branch in statistics called functional data analysis (FDA) (J. O. Ramsay & Silverman, 2005; J. Ramsay, Hooker, & Graves, 2009) that provides us with the mathematical tools required for such problems. FDA has found applications in different scientific fields such as climatology, biology and finance.

To start with, we assume an error model $y_j = x(t_j) + \varepsilon_j$, where $x(t_j)$ is a continuous, smooth analytical expression evaluated at time t_j and ε_j is some residual term. Our objective is to find that function $x(t)$, such that the residuals are minimized and the function is as smooth as possible. In other words, we have a regularized functional regression. We further assume that this model is additive and $x(t)$ is constructed from a linear combination of K simpler basis functions $\phi_k(t)$, such that $x(t_j) = \sum_k^K c_k \phi_k(t_j)$. We point the reader to (Eilers & Marx, 1996) for details on these functions. Then we can specify a least squares estimation problem

$$\mathrm{SSE}(\mathbf{c}) = \sum_j^n [y_j - \sum_k^K c_k \phi_k(t_j)]^2 = \sum_j^n [y_j - \phi^{'}(t_j)\mathbf{c}]^2, \tag{2}$$

where $\mathbf{c}$ is the vector of coefficients c_k and $\phi^{'}(t_j)$ is the jth row of the $n \times K$ matrix Φ, where $\Phi_{jk} = \phi_k(t_j)$. Next, we measure the roughness of the function $x(t)$ by introducing a penalty term, the *"integrated squared second derivative or total curvature"* (J. Ramsay et al., 2009)

$$\mathrm{PEN}_2(x) = \int [D^2 x(t)]^2 dt, \tag{3}$$

where the bounds of integration coincide with the start and end time of each trajectory (omitted here). Then, we can formulate the final expression (J. Ramsay et al., 2009) to be minimized as

$$\mathrm{PENSSE}_\lambda(\mathbf{c}) = \mathrm{SSE}(\mathbf{c}) + \lambda \mathrm{PEN}_2(x), \tag{4}$$

where λ is a hyperparameter that controls the effect of the penalty term. The higher the λ, the smoother the result. If $\lambda = 0$, it means that there is no regularization. This parameter is chosen by the *"generalized cross-validation criterion"* (J. Ramsay et al., 2009):

$$\mathrm{GCV}(\lambda) = \left(\frac{n}{n - df(\lambda)}\right)\left(\frac{\mathrm{SSE}}{n - df(\lambda)}\right). \tag{5}$$

The degree of freedom of the fit $df(\lambda)$ can be found if we define the *"symmetric roughness penalty matrix"* $R = \int \phi(t)\phi^{'}(t)dt$ of order K and let $df(\lambda) = \mathrm{trace}\left[\Phi(\Phi^{'}\Phi + \lambda \mathbf{R})^{-1}\Phi^{'}\right]$. The number of basis functions used is proportional to the duration of each trajectory, as in (Yang, Ozbay, Xie, Yang, Zuo, & Sha, 2021; Yang, Ozbay, Xie, Yang, & Zuo, 2021).

Penalizing false negatives

So far, we obtained error functions per trajectory for speed and azimuth. Clearly, we also need to specify the respective thresholds of what is considered anomalous. Please note that thresholds are set on the error functions x_v, x_θ and not directly on the metrics of interest v, θ. Let us assume for now that such thresholds x_t are known. If $x > x_t$, the respective observations of interest are marked as anomalous and are removed from the dataset. The missing values are then interpolated linearly. Finally, we obtain anomaly-free speed v^* and azimuth θ^*.

It is known that the majority of authors validate trajectory reconstructions based on longitudinal acceleration. Here, we will break away from this tradition and propose an alternative method based on speed and lateral acceleration. Lateral acceleration is very important in urban settings as it is related to phenomena of turning maneuvers and lane-changes. Therefore, we calculate the noise-free longitudinal acceleration $a^*_{lon} = dv^*/dt$ and lateral acceleration $a^*_{lat} = v^{*2} d\theta^*/ds^*$, where $s^* = \int v^* dt$. In absence of any ground truth, we will use insights gained from a study with instrumented vehicles (Bosetti, Da Lio, & Saroldi, 2014). The authors make some very interesting and strong arguments based on the theory of human-machine interaction and attempt to deduce its governing laws. They propose a *"modified Levinson criterion"* (Bosetti et al., 2014) that sets the feasible bounds on lateral acceleration as a function of speed, Figure 2. The *"Bosetti criterion"* is

$$a^{Bosetti}_{lat} = \frac{a_0}{\sqrt{(1-(v^*/v_0)^2)^2 + 2(v^*/v_0)^2}}, \tag{6}$$

where the intercept $a_0 = 5.22 m/s^2$ and the inflection point $v_0 = 14.84 m/s = 53.42 km/h$, according to the original paper. These parameters fit well urban traffic, but are inadequate for highways.

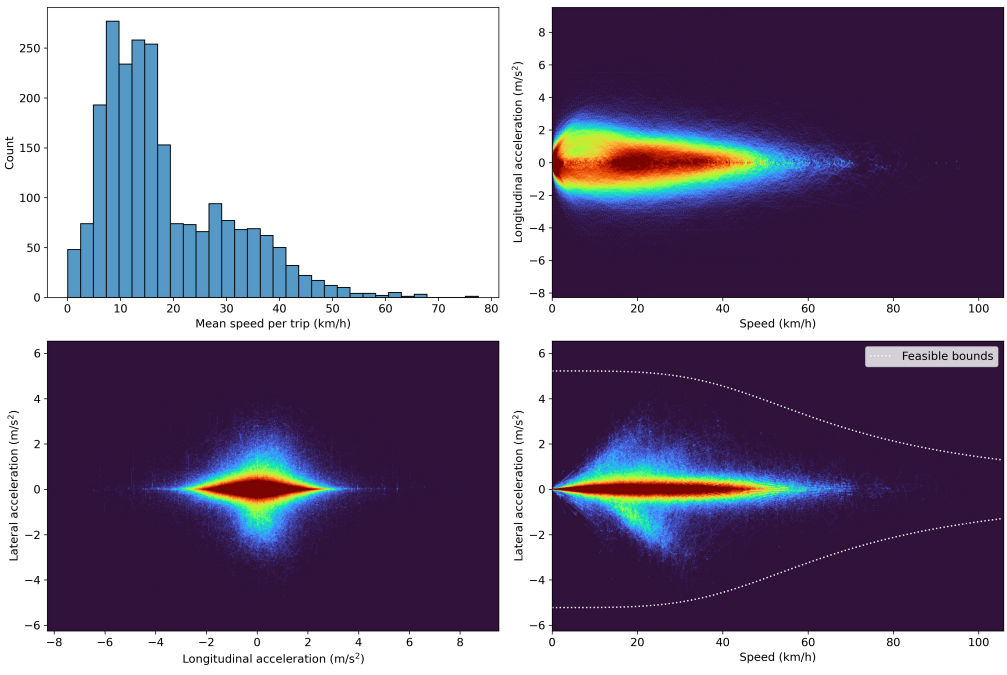
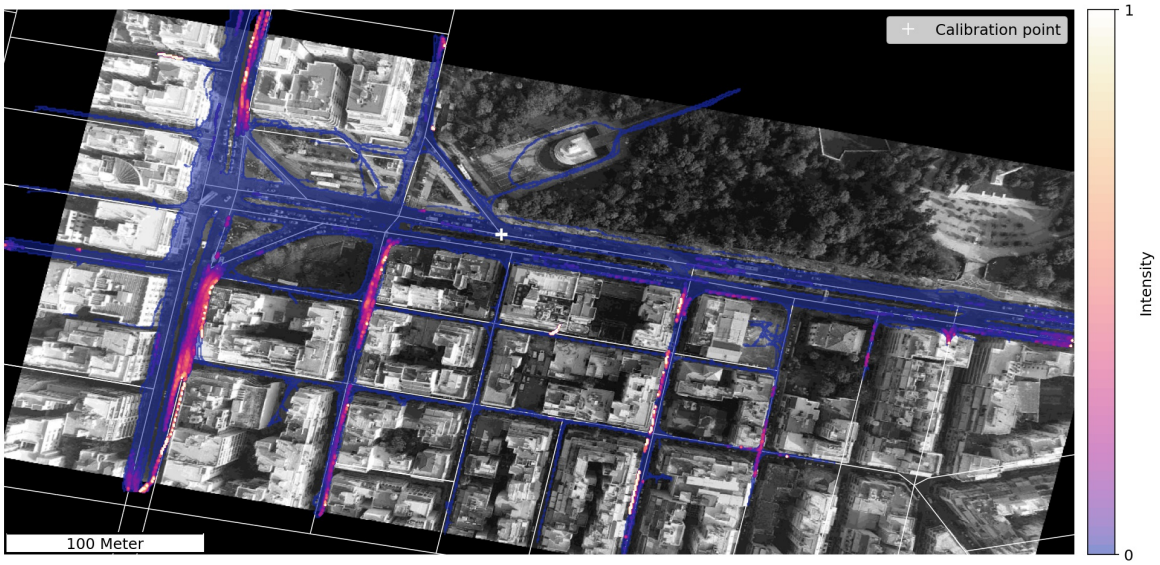


Figure 2: Speed, longitudinal and lateral acceleration for congested traffic, as shown by the histogram of average speed per trip. Color represents density of observations.

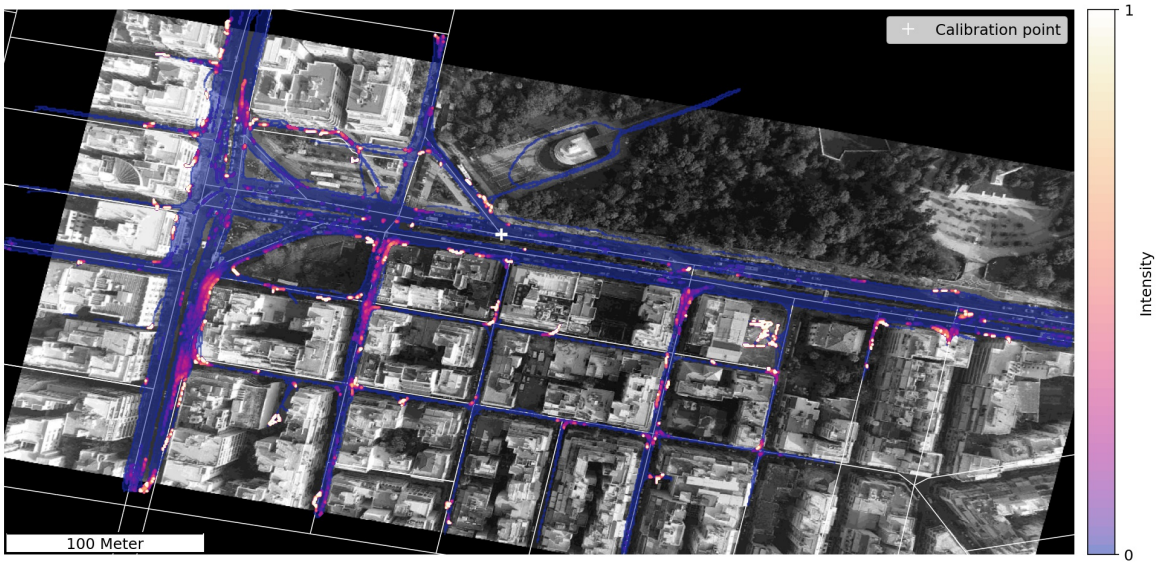
Every observation that is found outside of the feasible domain, as described by the Bosetti equation, is considered anomalous. This basically sets an upper bound for our reconstruction. Now we just need to specify the lower bound. Notice how close we are to the feasible domain only by removing anomalies and without even using smoothing.

Penalizing false positives

In principle, the anomaly-free distribution, as described in the previous section, can shrink arbitrarily in absence of a regularization mechanism that penalizes false positives. To this end, we introduce an elegant in its simplicity heuristic: a calibration point, as in Figure 3, is chosen such that the expected anomaly count in its immediate vicinity is close to zero. The calibration point is placed there where no visual restrictions or perspective distortions are expected. We set the vicinity radius $r = 3m$. It is also worth mentioning that the selected location has been strategically placed on the main arterial in order to control as much of the traffic as possible. Both speed and azimuth anomalies are included in the count.



(a) Speed anomaly heatmap. The main arterial has few anomalous observations.



(b) Azimuth anomaly heatmap. Some anomalies are due to parking maneuvers.

Figure 3: Penalizing false positives with a calibration point.

For illustrative purposes, we discretize space in small enough bins and define the intensity as the normalized anomaly count per bin, excluding static points. Both figures support our hypothesis of spatial autocorrelation as the most noisy areas are parts of the network that the drone had limited visibility. A less obvious finding is related to pronounced anomalies near the edges that can be safely attributed to perspective distortions.

Determining the optimal parameters

By now we have gained a qualitative understanding of the objective and we proceed to formulate a two-dimensional optimization problem by minimizing the following loss function

$$L(x_t^v, x_t^\theta) = \sum_m^M \{|a_{lat}^*(m)| - a_{lat}^{Bosetti}(m)\} + \text{Card}\{\mathbf{p}|\text{d}(\mathbf{p},\mathbf{q}) < r\}, \tag{7}$$

where x_t^v, x_t^θ are the speed and azimuth error thresholds, M is the total number of undetected anomalous observations (false negatives) and the last term is the cardinality of the set of misclassified points $\mathbf{p}$ that fall within a distance r from the calibration point $\mathbf{q}$ (false positives). This loss function is the combination of one floating point term and one integer term. The two terms balance because a strong decrease in false negatives will drastically increase the false positives and will ultimately increase the value of the objective function. It turns out that, for the problem under consideration, our objective function has a single, well defined global optimum (minimum) and it is convex, Figure 4. Please be aware that this statement holds true for the particular filter specification with $\Omega_c = 0.9$ Hz. Higher cut-off frequencies perform poorly and lower cut-off frequencies produce multiple minima.

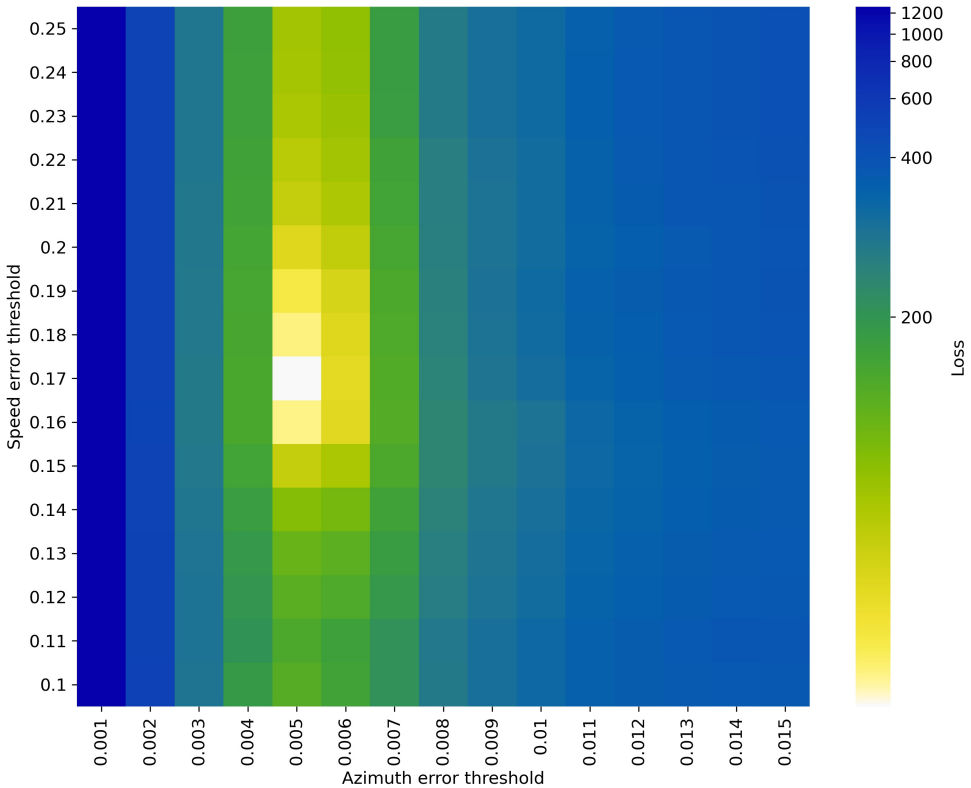


Figure 4: Evaluations of the objective function for different values of speed and azimuth error thresholds.

The optimization is done by multi-threaded, direct evaluation on a 15×15 grid and the search ranges have been identified by inspection. Notice that the two thresholds differ by one order of magnitude. Inside the optimum grid cell, polishing in the form of local search finds the exact threshold parameters.

3. RESULTS AND DISCUSSION

With the optimal parameters in place, we can quantify the errors in the dataset. These are summarized in Table 1. We distinguish between contamination rates and anomaly rates. A trajectory is contaminated if any of its observations is anomalous, Figure 5. Static points are excluded in anomaly counts.

Table 1: Result of anomaly detection for speed and azimuth

			Anomaly statistics	
Error metric	Speed	Azimuth	Both	Union
Portion of contaminated vehicle trajectories.	19.62%	25.61%	11.85%	33.39%
Portion of anomalous data, excluding staypoints.	0.81%	0.93%	0.24%	1.5%

As mentioned in the beginning, we make a clear distinction between anomalies and white-noise. Previous research has focused on the second type of errors with extensive use of smoothing. In reality, most of the contribution to the noise comes from the first type of errors. Nevertheless, the residual white-noise should be also treated. This is especially relevant for the accelerations because errors are amplified by derivation, Figure 6.

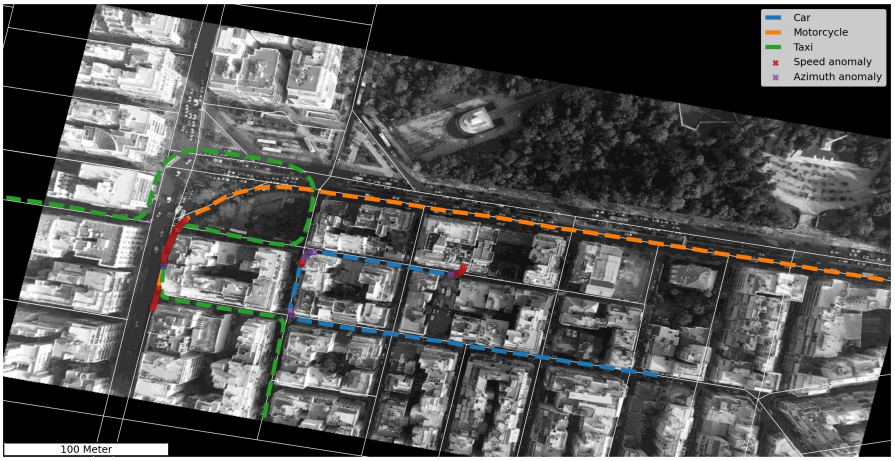


Figure 5: Anomaly detection for 3 different trajectories of various vehicle types.

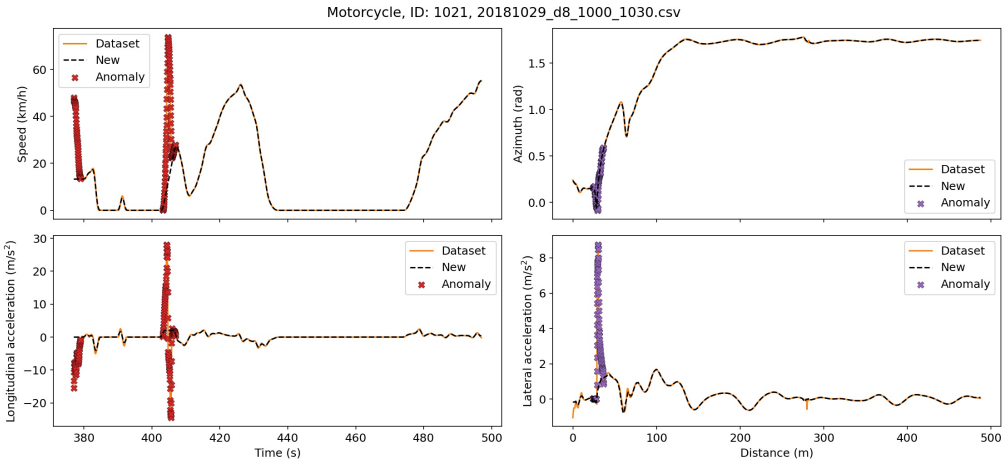


Figure 6: Anomaly detection and smoothing for the motorcycle trajectory.

For the elimination of white-noise, we use a first order Savitzky-Golay (SG) filter with a window length equal to the frame rate (25 fps). This is equivalent to a centralized moving average. Centralized moving averages are a good choice for offline data as they do not introduce delay effects, (Haslwanter, 2021). The SG filter is applied on the anomaly-free data. From the level of a single vehicle trajectory to the overall distribution of lateral and longitudinal accelerations in the dataset, Figure 7, before and after reconstruction, we can see how our treatment reduces the noise without introducing noticeable systematic oversmoothing.

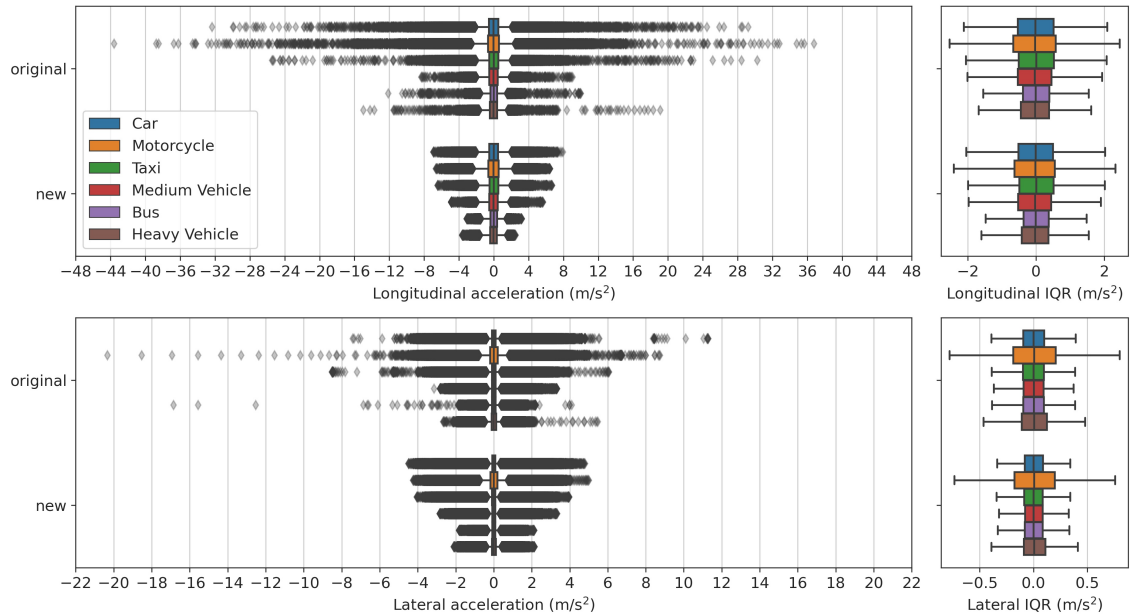


Figure 7: Distribution of longitudinal and lateral accelerations before and after our reconstruction. The respective interquantile ranges remain mostly intact.

At this point, we reiterate that the above distributions exclude static points, known also as stay-points. This is very important and it has been recently shown by (Paipuri, Barmpounakis, Geroliminis, & Leclercq, 2021) that static and moving vehicles are completely different phenomena and should be treated separately. The emergence of detailed data on urban traffic has triggered new interest in the empirical study of this distinction.

On another note, some authors put emphasis on the importance of internal and/or external consistency of vehicle trajectory reconstructions (Montanino & Punzo, 2015; Dong et al., 2021). This idea is interesting, but not without significant shortcomings, see (Coifman & Li, 2017) for a detailed criticism. We should add here, that especially the concept of external, or platoon consistency fails to account for the spatial autocorrelation of the errors among neighboring trajectories.

4. CONCLUSIONS

In summary, we presented a novel detection method for speed and for azimuth anomalies. We managed to recover, as faithfully as possible, the underlying distributions of longitudinal and lateral accelerations for six different modes of transport without hand-crafted rules. Our method can be used for quantifying autocorrelated errors or, inversely, for identifying network segments that are devoid of errors and can thus be used directly for validation of traffic flow models or for traffic safety analytics. In future research, we would like to exploit the denoised data for studying the characteristics of lateral interactions in the urban environment with emphasis in motorcycles.

ACKNOWLEDGMENT

Special thanks goes to Dr. Gustav Nilsson for his valuable elaboration on the Butterworth filter. His suggestions have been crucial to the development of this work. This research was partially funded by Swiss National Science Foundation (SNSF) grant (200021_188590) "pNEUMA: On the new era of urban traffic models with massive empirical data from aerial footage".

REFERENCES

Barmpounakis, E., & Geroliminis, N. (2020). On the new era of urban traffic monitoring with massive drone data: The pneuma large-scale field experiment. *Transportation research part C: emerging technologies*, *111*, 50–71.

Bosetti, P., Da Lio, M., & Saroldi, A. (2014, Jun 01). On the human control of vehicles: an experimental study of acceleration. *European Transport Research Review*, *6*(2), 157-170.

Coifman, B., & Li, L. (2017). A critical evaluation of the next generation simulation (NGSIM) vehicle trajectory dataset. *Transportation Research Part B: Methodological*, *105*, 362-377.

Dong, S., Zhou, Y., Chen, T., Li, S., Gao, Q., & Ran, B. (2021). An integrated empirical mode decomposition and butterworth filter based vehicle trajectory reconstruction method. *Physica A: Statistical Mechanics and its Applications*, *583*, 126295.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statistical Science*, *11*(2), 89 – 121.

Haslwanter, T. (2021). *Hands-on signal analysis with python: An introduction* (1st ed.). Cham, Switzerland: Springer Nature.

Montanino, M., & Punzo, V. (2015). Trajectory data reconstruction and simulation-based validation against macroscopic traffic patterns. *Transportation Research Part B: Methodological*, *80*, 82-106.

Paipuri, M., Barmpounakis, E., Geroliminis, N., & Leclercq, L. (2021). Empirical observations of multi-modal network-level models: Insights from the pneuma experiment. *Transportation Research Part C: Emerging Technologies*, *131*, 103300.

Ramsay, J., Hooker, G., & Graves, S. (2009). *Functional data analysis with R and MATLAB* (2009th ed.). New York, NY: Springer.

Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis* (2nd ed.). New York, NY: Springer.

Taylor, F. (2011). *Digital filters: Principles and applications with MATLAB*. New York, NY: Wiley-IEEE Press.

Vlahogianni, E. I., & Barmpounakis, E. N. (2017). Driving analytics using smartphones: Algorithms, comparisons and challenges. *Transportation Research Part C: Emerging Technologies*, *79*, 196–206.

Yang, D., Ozbay, K., Xie, K., Yang, H., & Zuo, F. (2021). A functional approach for characterizing safety risk of signalized intersections at the movement level: An exploratory analysis. *Accident Analysis & Prevention*, *163*, 106446.

Yang, D., Ozbay, K., Xie, K., Yang, H., Zuo, F., & Sha, D. (2021). Proactive safety monitoring: A functional approach to detect safety-related anomalies using unmanned aerial vehicle video data. *Transportation Research Part C: Emerging Technologies*, *127*, 103130.